DOCUMENT RESUME

ED 466 957 SE 066 503

AUTHOR Abraham, Paul

TITLE A Microscopic Look at Assessment: Dropping a Lowest Test Score versus Allowing a Retake Test.

PUB DATE 2000-00-00

NOTE 9p.; In: Proceedings of the Regional Campus Mathematics Conference (Canton, OH, May 19-21, 1999). AURCO stands for Association for University Regional Campuses of Ohio.

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

JOURNAL CIT AURCO Mathematics Journal; v1 2000

EDRS PRICE EDRS Price MF01/PC01 Plus Postage.

DESCRIPTORS *Grades (Scholastic); *Evaluation; Grading; Higher Education; *Mathematics Education; *Testing

ABSTRACT

This paper describes various methods of calculating final grades for mathematics courses. Specifically, a comparison between dropping the lowest test score from the semester and allowing a retake of one test for an improved grade is made. It is concluded that a retake test policy is preferable to a dropped test policy because students must work to improve their grades; student final exam performance was higher with retakes; and students supported the retake policy in written comments. (MM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 466 957

A Microscopic Look At Assessment: Dropping A Lowest Test Score Versus Allowing A Retake Test

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

P. Abraham

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paul Abraham

Kent State University Stark

INTRODUCTION

In more than half of my full-time teaching experience, I have dropped the lowest test score in calculating semester grades in many of my low-level mathematics courses. Intuitively, a dropped test score seemed reasonable since I tested frequently; hence, I had several scores to form a representative test average after dropping the lowest score. Professionals with more expertise in assessment have also considered dropping the lowest score as a reasonable compromise. In discussing a unit-based approach for grading (which stresses importance of all units) versus a developmental approach (which stresses importance of skills at the end), Walvoord and Anderson [7] suggest that "a middle ground is to count all tests and assignments heavily but to allow the student to drop his or her lowest grade." On the other hand, Shalom [5] shows the anti-intuitive result that dropping the lowest (or highest) score can commonly increase standard deviation in a student's test scores.

Furthermore, since dropping a test score improves all student course averages without any evidence of added student effort, concern for grade inflation and student learning arose in my mind at the outset. To address both concerns, I relied on a comprehensive final exam weighted around 25% in semester grades to provide a double check of student understanding of all course materials. Yet, in my experience student scores on comprehensive final exams in dropped test sections have traditionally been low, averaging about 64% in the sections included in this study. Thus, I wondered if the final exam really did encourage students to restudy dropped materials. That issue was especially poignant since generally the last test before the final exam was most often dropped yet covered important and the least practiced material for future math classes, e.g., logarithm and exponential functions and equations in college algebra for calculus students.

Keeping these issues in mind, rather than dropping the lowest score I recently offered students in several of my classes the opportunity for one retake test score to replace a lower corresponding test score. In forcing students to be active participants in improving their course averages, the retake policy at least alleviated some of my concerns about student learning and grade inflation. Section 1 will further describe the retake policy and procedure.

Besides my ease of mind, other potential benefits of the retake policy surfaced while collecting data for this study. The data are displayed in **Tables 1**, **2** and **3** in Sections 2 and 3 without any

SE006583

BEST COPY AVAILABLE

assertion of statistical significance. **Table 1** illustrates the view that the retake policy predictably yielded lower course averages than the dropped test policy would have produced in the retake test sections, thus reducing grade inflation while still giving motivated students a chance to recover from a low test score. **Table 2** demonstrates the surprising amount of increase in student averages caused by dropping the lowest test score in dropped test sections. **Table 3** suggests the retake policy was less lenient in terms of assigned grades and student GPAs, with little effect on my student evaluations. The data also suggest that students in retake sections tended to predict lower but more realistic semester grades. Finally, Sections 4 and 5 supply some encouraging evidence that students have appreciated and benefited from the retake policy.

1. ADMINISTRATION OF RETAKES

In each retake test section, I informed the students orally at the beginning of the semester and with a statement in my syllabus that they would be offered the opportunity to have one retake test score be used in computing their course average. Some basic rules were that there would be only one retake test per original test, a retake test score lower than the original test score would be not be counted, and there would be no retake opportunity for the final exam. The policy also permitted students to retake more than one test in case of two or more low original test scores. However, I explained only one retake test score—the one that represented the greatest gain in points—would be used as a replacement for the corresponding test score.

I announced test and retake dates in my syllabus so students could plan ahead accordingly. I generally scheduled around 10 days between the test and retake dates to allow time for me to grade and return the tests and retakers about a week to restudy test materials and resolve any questions from the original test or homework assignments. Students normally completed retakes on the designated days on a walk-in basis at the testing center. I typically chose retake dates to be consecutive Mondays and Tuesdays, the days that the testing center was open for extended hours. Since students could not leave the testing center with the retake test or any work paper, potential cheating was limited to what problems a student retaking a test first could memorize and restate to a coconspirator. In general, I found myself less concerned with the potential cheating on retakes than other alternate forms of assessment like group work or take-home materials. At times, though, I did send multiple versions of a retake test to the testing center.

In constructing retakes, I allowed about 60% of the problems to be very similar to problems on the original test. Roughly the next 20% of the retake test problems were somewhat similar to original problems in which, for example, a retake and original test problem would have the same overall goal but involve slightly different strategies. The remaining retake problems were somewhat different. For example, to discourage memorization, word problems on the retake test generally involved different setups than word problems on the original test.

2. A COMPARISON OF COURSE AVERAGES

Table 1 summarizes course averages in the four sections of intuitive calculus (Math 11012) and two sections of college algebra (Math 11011) for which I offered students the retake test policy. It incorporates data on the approximately 150 fullterm students (took all tests and the final), about 40 in college algebra, and 110 in intuitive calculus. Besides the section retake

average μ_{RA}, the table contains three other hypothetical section averages: (1) section drop average (μ_{DA})—having a dropped test but no retake test, (2) section straight average (μ_{SA})—having no dropped test and no retake test, and (3) section drop/straight average ($\mu_{D/SA}$)—averaging drop and straight average. To avoid any distortion in the averages due to the varying number of total points involved in each grading scheme, the percentage of each grading factor was held the same in computing RA, DA, SA, and D/SA. The small $\mu_{RA-D/S}$ values imply that D/SA provides an excellent estimate (on the average) for RA.

Table 1: Student Course Average Information for Retake Test Sections

Section	Full-term Students	Percent Retakers	μ_{RA}	μ_{DA}	μ_{SA}	$\mu_{D/SA}$	$\mu_{RA-D/SA}$
Intuitive Calculus (F97, morning)	36	72.22	0.80998294	0.82009	0.794867	0.807479	0.002504
Intuitive Calculus (F97, evening)	25	30.77	0.78956667	0.80237	0.7758	0.795967	-0.0064
Intuitive Calculus (F98, morning)	22	63.64	0.79	0.79922	0.772845	0.786031	0.003928
Intuitive Calculus (F98, evening)	24	62.50	0.73418822	0.75065	0.717306	0.733979	0.00021
College Algebra (S99, early afternoon	24	45.83	0.73141061	0.74337	0.717987	0.730679	0.000732
College Algebra (S99, late afternoon)	17	41.18	0.81797484	0.82868	0.809764	0.819224	-0.001249
Weighted Average		54.60	0.779462	0.79128	0.765039	0.779322	0.000121

Next, **Table 2** displays course average information for four sections of college algebra, the courses I most recently graded using the dropped test policy. The table incorporates data on the approximately 80 fullterm students in those sections. As in **Table 1**, μ_{DA} is compared with hypothetical μ_{SA} and $\mu_{D/SA}$.

Table 2: Student Course Average Information for Dropped Test Sections

Section	Full-term Students	μ_{DA}	μ_{SA}	$\mu_{D/SA}$
College Algebra (S97, afternoon)	22	0.79511	0.75358	0.77435
College Algebra (S97, late afternoon	24	0.79942	0.7639	0.78166
College Algebra (S98, afternoon)	21	0.76554	0.72571	0.74563
College Algebra (S98, evening)	15	0.80284	0.76926	0.78605
Weighted Average		0.79021	0.75233	0.77127

BEST COPY AVAILABLE

Finally, the average μ_{DA} value in **Table 1** is about 2.5% higher than the average μ_{SA} value, whereas the average μ_{RA} value is only about 1.4% higher the average μ_{SA} value. If the μ_{SA} and μ_{DA} values in **Tables 1** and **2** were combined, the average μ_{DA} value would surpass the average μ_{SA} value by 3.0%. These figures demonstrate that the dropped test policy produced about twice as much grade inflation (using SA as the standard) than the retake policy.

3. GPA AND STUDENT EVALUATION SCORES

The first three columns of **Table 3** below supply the section average grade received (μ_{GR}), average student GPA (μ_{GPA}), and leniency score (simply μ_{GPA-GR} as discussed in [6]) for the ten classes in this study. The GPA information obtained from the Student Information System is based on the standard 4.0 scale and is cumulative as of the end of the semester the particular section was completed. The GPA information is somewhat crude since it includes all students per official class roster, including those who withdrew from or just stopped attending the course. **Table 3** indicates that the retake sections produced an average section leniency score 17.9% $((-.33 - -.28)/(-.28))$ more negative than the dropped test sections despite exhibiting a slightly higher average section GPA. The latter, of course, was probably due to the fact that only 54.6% of all fullterm students worked to improve their grades with a retake.

The remaining columns are items collected from the standard KSU student evaluations given to each of my classes. The data includes my section and the corresponding campus scores on the average grade expected by the student (μ_{GE}), Question 20 "overall the instructor's teaching was . . ." (μ_{Q20}), and Question 16 "methods of assigning grades was fair" (μ_{Q16}). Student evaluation scores range from 1 to 5, where 1 is considered best.

Table 3. Average Grade and Student Evaluation Information for Retake Sections

	Section μ_{GR}	Section μ_{GPA}	Section Leniency μ_{GR-GPA}	Section μ_{GE}	Campus μ_{GE}	Section μ_{Q20}	Campus μ_{Q20}	Section μ_{Q16}	Campus μ_{Q16}
Retake Test Sections									
Intuitive Calculus (F97, morning)	2.68	2.84	-0.16	1.83	1.84	1.2	1.98	1.4	1.88
Intuitive Calculus (F97, evening)	2.65	2.64	0.01	1.65	1.84	1.3	1.98	1.25	1.88
Intuitive Calculus (F98, morning)	2.35	2.77	-0.42	2.24	1.88	1.86	1.96	1.76	1.87
Intuitive Calculus (F98, evening)	2.04	2.72	-0.68	2.68	1.88	1.6	1.96	1.4	1.87
College Algebra (S99, early afternoon)	1.88	2.45	-0.57	2.26	1.82	1.43	1.88	1.5	1.8
College Algebra (S99, late afternoon)	2.47	2.74	-0.27	2.07	1.82	1.4	1.88	1.33	1.8
Weighted Average	2.37	2.7	-0.33	2.1	1.85	1.44	1.95	1.44	1.85
Dropped Test Sections									
College Algebra (S97, afternoon)	2.55	2.65	-0.1	1.69	1.84	1.35	1.94	1.18	1.85
College Algebra (S97, late afternoon)	2.13	2.45	-0.32	1.77	1.84	1.38	1.94	1.31	1.85
College Algebra (S98, morning)	2.09	2.58	-0.49	2	1.79	1.82	1.95	1.41	1.85
College Algebra (S98, evening)	2.22	2.41	-0.19	1.5	1.79	1.3	1.95	1.4	1.85

The expected grade information in **Table 3** supports the conclusion that students expect grade inflation. For example, in [3] the author shows that despite students having a reasonably clear idea that "average" work corresponds to a C grade, most students who rated their work as average near the end of the semester expected to receive an A or B in the course. Similarly, in most of the sections in this study, the students expected significantly higher grades than received, despite receiving updates on their course averages after each test and a precise grading scale on the course syllabus. The retake students, however, were much more accurate in predicting grades than were the dropped-test students in predicting grades. The average predicted grade value in the retake sections was 2.10, which corresponds to 2.90 on the 4.0 scale, whereas the average predicted grade value in the dropped test sections was 1.76, which translates to a highly significant 3.24 on the 4.0 scale.

Although I, as many other educators, tend to question the general validity of the numerical student evaluations, I wondered if the less lenient retake policy versus the more lenient dropped test grading method would produce any obvious difference in student evaluation scores. Student evaluation data in **Table 3** suggest little difference in the two measures which seemed relevant: my overall rating value μ_{Q20} and my rating on fairness of grading methods value μ_{Q16}. An interesting

BEST COPY AVAILABLE

observation to the contrary is given by Johnson and Beck [2] in which instructors applied a strict grading scale in some educational psychology sections or a lenient scale in other sections. One of the findings was that although low SAT students scored higher on tests in the strict grading scale sections and moderate and high SAT students scored about the same, students in stricter scale sections gave the instructors worse evaluations.

4. A BRIEF COMPARISON OF STUDENT PERFORMANCE ON FINAL EXAMS

So far the data that I have collected seem to indicate that student performance on final exams has improved in two ways. First, students in retake sections have scored considerably better on final exams than students in dropped test sections, averaging about 73% correct in retake sections versus 64% correct in dropped test sections. It is difficult, though, to conclude what part of the difference is due purely to the retake policy since other changes occurred in the time span of this study, for instance, the emphasis of the final, format of the final, etc.

However, a more focused look at final exams revealed a second interesting, and perhaps more significant, form of improvement. Since the last test before the final exam was by far the most popular test to drop or retake in all sections, I examined student performance on final exam questions corresponding to the last test. I had an idea that most students who dropped the last test conserved energy by not restudying the last test material for the final exam. So, I examined final exam papers of those students who dropped the last test in my spring 1997 college algebra courses and those who completed a retake on the last test in my fall 1998 intuitive calculus courses. I chose those two courses in those two semesters since the format and comprehensive nature and weighting of the final exam in student grades were very similar. Indeed, my hunch about low student performance was unfortunately correct. The performance on final exam questions corresponding to the last test was incredibly large: on the average 66% correct in the retake sections versus 40% correct in the dropped test sections.

5. WRITTEN STUDENTS COMMENTS ABOUT THE RETAKE POLICY

Since I received no prior feedback about the retake policy from the written segment of KSU student evaluation forms, I asked my spring 1999 college algebra students to complete a separate survey form. I passed out the form near the end of the semester after the last retake test date had passed. To insure secrecy of student responses until my grades were submitted, I followed the same procedure in handling the surveys as student evaluation forms.

Question 1 asked if the students understood the retake policy. Most agreed they did.

Question 2 asked, "Did the retake policy help you learn materials better?" Some responses going beyond "yes" were:

- "It (the retake) gave me time to look over the last test and see my mistakes."
- "Because I got a better grade (on the retake)."
- "After taking the first test and getting a bad grade, knowing that I had another chance made me study harder."
- "It gave me confidence in my work by knowing that I could have a 2nd chance to do better if I didn't do too well the first time."

- "I had to look deeper into the material."

Question 3 asked, "If you didn't use a retake, why not?" Many responded that they were satisfied with the original test scores, while several responded they did not have time or the retake dates were inconvenient. Some other comments were:

- "I chose not to take a retake because I didn't think I would do all that much better."
- "I forgot, plus, when I did remember, I knew that the (retake) tests would be harder."
- "I heard they [retakes] were harder than the first test. Also, I figured I did not want to go crazy over a test twice. I figured I did what I could on the first one."
- "To get a meaningful measure of my knowledge, I feel that I should go with a test as taken. I could have used the points, but I wanted to know how I really did. I had time to study, and, if I didn't, I feel I should pay the price."

Question 4 asked for any additional comments about the retake policy. All comments endorsed or appreciated a retake test opportunity. Some of comments proposed different retake schemes:

- "Let us take each one over once!"
- "I think the retake is a good idea. As long as the material is learned by the end of the semester, then the teacher has done his job."
- "I think maybe you should let them, the students, take all the tests and then retake their lowest before the final; the final should not be included; that would be less stressful."

The first two comments relate to grading practices already used in other disciplines. For example, the first comment compares to "rewrite days" for papers as discussed by writing instructors Logan [4] and Flores [1] respectively. The second comment reflects mastery learning, as discussed for example by marketing instructor Webster [8], in which the student must retest repeatedly until achieving a specified level of competency. Finally, the last comment and responses to Question 3 indicating scheduling problems with the retakes correspond to the possibility that I have considered on occasion of using the final exam period as the single time for students to retake a test of their choice. However, at this stage, I still cling to the importance of a comprehensive final exam.

CONCLUSION

Although preliminary, the data presented in this paper show that a retake test policy is preferable to a dropped test policy for several reasons. First, students must work to improve their grades with a retake test, thus lowering grade inflation while still giving motivated students the opportunity to improve their course averages. Second, student final exam performance was somewhat higher in the retake sections. Yet the retake sections exhibited less grade inflation without any significant negative effect on my student evaluation scores. Finally, in their written comments, students supported the retake policy. About the only negative aspect of the retake policy is the increased workload for the instructor.

ACKNOWLEDGMENT

I would like to thank Austin Melton, KSU Mathematics and Computer Science Chair, for sending my paper to the referees and the referees for their valuable suggestions in revising this paper.

REFERENCES

[1] Flores, J. (1997). "Grading from a 'No Curve' Pedagogy." In Libby Allison, Lizbeth Bryant, and Mareen Hourigan (Eds.) *Grading in the Post-Process Classroom*, Chapter 8, (pp. 120-124). Boyton/Cook Publishers.

[2] Johnson, B. and Beck, H. (1998). "Strict and Lenient Grading Scales: How Do They Affect the Performance of College Students with High and Low SAT Scores." *Teaching of Psychology*, 15 (3), 127-131.

[3] Landrum, R. (1999). "Student Expectations of Grade Inflation." *Journal of Research and Development in Education*, 32 (2) 124-128.

[4] Logan, L. (1995). "Making Writing Real: 'Rewrite Days' and Other Empowerments." *College Teaching*, 43, 132-133.

[5] Shalom, S. (1995). "Dropping Scores." *Mathematics Magazine*, 68 (3), 220-223.

[6] Tkacz, S. (1999). "Factors Affecting Student Evaluations of Faculty." *OATYC Journal*, 24 , 13-20.

[7] Walvoord, B. and Anderson, V. (1998) *Effective Grading*. San Francisco: Jossey-Bass Inc., 98-99.

[8] Webster, M. (1994). "Try, Try Again." *Vocational Educational Journal*, 69, 30-32.

BIOGRAPHICAL AND ADDRESS INFORMATION

Paul Abraham received his Ph.D. from Kent State University in mathematics in 1993 under the supervision of Joe Diestel. His research interests include Banach spaces, martingale theory, and teaching-related issues such as course assessment and the use of technology in the classroom. He joined the KSU Stark faculty in 1996 after teaching for four years at the College of the Ozarks in Missouri, where he first became interested in course assessment.

Kent State University Stark
6000 Frank Ave NW
Canton OH 44720
pabraham@stark.kent.edu

SE016503

ERIC

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: (article) A Microscopic Look: Dropping a Lowest Test Score Versus Allowing a Retake Test	
Author(s): Paul Abraham	
Corporate Source:	Publication Date: Aug. 2000

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

✓

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: Paul Abraham	Printed Name/Position/Title: Paul Abraham, Assistant Professor of Mathematics	
Organization/Address: Kent State Stark Campus, 6000 Frank Ave NW, Canton OH 44720	Telephone: 330 244 3404	FAX: 330 494-6121
	E-Mail Address: pabraham@stark.kent.edu	Date: February 15, 2002